



(12) 发明专利申请

(10) 申请公布号 CN 115239715 A

(43) 申请公布日 2022. 10. 25

(21) 申请号 202211154439.0

G06V 10/44 (2022.01)

(22) 申请日 2022.09.22

G06V 10/764 (2022.01)

(71) 申请人 中南大学

G06V 10/774 (2022.01)

地址 410000 湖南省长沙市岳麓区麓山南路932号

G06V 10/82 (2022.01)

G06V 20/69 (2022.01)

(72) 发明人 吴瑞文 黄伟红 黄佳 于永福
刘冠宇 李靖 高武强

(74) 专利代理机构 广州嘉权专利商标事务所有
限公司 44205

专利代理师 王本晋

(51) Int. Cl.

G06T 7/00 (2017.01)

G06N 3/04 (2006.01)

G06V 10/20 (2022.01)

G06V 10/26 (2022.01)

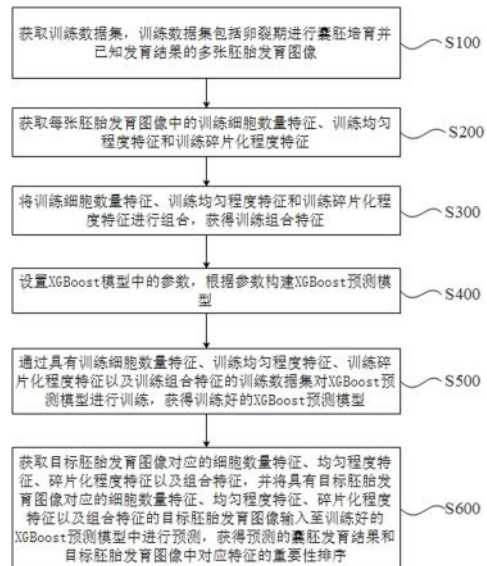
权利要求书4页 说明书18页 附图4页

(54) 发明名称

一种囊胚发育结果预测方法、系统、设备及存储介质

(57) 摘要

本发明公开了一种囊胚发育结果预测方法、系统、设备及存储介质,本方法通过获取训练数据集;获取每张胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征;将细胞数量特征、均匀程度特征和碎片化程度特征进行组合,获得训练组合特征;构建XGBoost预测模型;通过训练数据集对XGBoost预测模型进行训练,获得训练好的XGBoost预测模型;将具有细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。本发明能够提高预测囊胚发育结果的准确率。



1. 一种囊胚发育结果预测方法,其特征在于,所述囊胚发育结果预测方法包括:

获取训练数据集,所述训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像;

获取每张所述胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征;

将所述训练细胞数量特征、所述训练均匀程度特征和所述训练碎片化程度特征进行组合,获得训练组合特征;

设置XGBoost模型中的参数,根据所述参数构建XGBoost预测模型;

通过具有所述训练细胞数量特征、所述训练均匀程度特征、所述训练碎片化程度特征以及所述训练组合特征的训练数据集对所述XGBoost预测模型进行训练,获得训练好的XGBoost预测模型;

获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,并将具有所述目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序。

2. 根据权利要求1所述的囊胚发育结果预测方法,其特征在于,所述获取训练数据集,包括:

获取在卵裂期的第一时间内进行囊胚培育并已知发育结果的多张第一胚胎发育图像,并将所有所述第一胚胎发育图像作为第一训练数据集;

获取在卵裂期的第二时间内进行囊胚培育并已知发育结果的多张第二胚胎发育图像,并将所有所述第二胚胎发育图像作为第二训练数据集;其中,所述第二时间与所述第一时间不同;

根据所述第一训练数据集和所述第二训练数据集生成训练数据集。

3. 根据权利要求2所述的囊胚发育结果预测方法,其特征在于,所述目标胚胎发育图像包括第一目标胚胎发育图像和第二目标胚胎发育图像,所述第一目标胚胎发育图像和所述第二目标胚胎发育图像为目标胚胎对应的不同时间采集的胚胎发育图像,所述获取目标胚胎发育图像对应的细胞数量特征,包括:

采用Sobel算子和直方图均衡化对所述第一训练数据集中的每张所述第一胚胎发育图像进行图像增强处理,获得增强处理后的第一训练数据集;

在所述增强处理后的第一训练数据集中随机选取第一预设值的第一胚胎发育样本图像;

通过所述增强处理后的第一训练数据集对第一mask-RCNN实例分割模型进行分割训练,每次训练前对每张所述第一胚胎发育样本图像翻转一次,直到所述第一mask-RCNN实例分割模型中的损失值收敛时结束训练,获得训练好的第一mask-RCNN实例分割模型;

通过所述训练好的第一mask-RCNN实例分割模型对所述第一目标胚胎发育图像进行实例分割,获得所述第一目标胚胎发育图像中的第一细胞数量特征和第一分割图像;

采用Sobel算子和直方图均衡化对所述第二训练数据集中的每张所述第二胚胎发育图像进行图像增强处理,获得增强处理后的第二训练数据集;

在所述增强处理后的第二训练数据集中随机选取第二预设值的第二胚胎发育样本图

像；

通过所述增强处理后的第二训练数据集对第二mask-RCNN实例分割模型进行分割训练,每次训练前对每张所述第一胚胎发育样本图像翻转一次,直到所述第二mask-RCNN实例分割模型中的损失值收敛时结束训练,获得训练好的第二mask-RCNN实例分割模型；

通过所述训练好的第二mask-RCNN实例分割模型对所述第二目标胚胎发育图像进行实例分割,获得所述第二目标胚胎发育图像中的第二细胞数量特征和第二分割图像。

4. 根据权利要求3所述的囊胚发育结果预测方法,其特征在于,所述获取目标胚胎发育图像对应的均匀程度特征,包括:

获取所述第一训练数据集中标记为均匀、不均匀和严重融合的三类第一胚胎发育图像,并获取三类所述第一胚胎发育图像中对应的每个细胞的第一轮廓；

根据所述每个细胞的第一轮廓,分割获得三类所述第一胚胎发育图像中的多个细胞；

计算三类所述第一胚胎发育图像中对应的多个细胞中随机一个细胞的所有像素值,并计算所述细胞的所有像素值的第一中位数；

根据所述第一中位数,从三类所述第一胚胎发育图像中对应的所述第一分割图像中选取均匀、不均匀和严重融合对应的第一像素值范围；

计算所述第一目标胚胎发育图像对应的第一分割图像中所有细胞对应的第一目标中位数；

根据所述第一像素值范围和所述第一目标中位数,获取所述第一目标胚胎发育图像中的第一均匀程度特征；

获取所述第二训练数据集中标记为均匀、不均匀和严重融合的三类第二胚胎发育图像,并获取三类所述第二胚胎发育图像中对应的每个细胞的第二轮廓；

根据所述每个细胞的第二轮廓,分割获得三类所述第二胚胎发育图像中的多个细胞；

计算三类所述第二胚胎发育图像中对应的多个细胞中随机一个细胞的所有像素值,并计算所述细胞的所有像素值的第二中位数；

根据所述第二中位数,从三类所述第二胚胎发育图像中对应的所述第二分割图像中选取均匀、不均匀和严重融合对应的第二像素值范围；

计算所述第二目标胚胎发育图像对应的第二分割图像中所有细胞对应的第二目标中位数；

根据所述第二像素值范围和所述第二目标中位数,获取所述第二目标胚胎发育图像中的第二均匀程度特征。

5. 根据权利要求4所述的囊胚发育结果预测方法,其特征在于,所述获取目标胚胎发育图像对应的碎片化程度特征,包括:

获取所述第一目标胚胎发育图像中的所有第一胚胎像素值；

对所有所述第一胚胎像素值进行遍历,获取第一极大像素值；

根据所述第一极大像素值的个数划分所述第一目标胚胎发育图像中的第一碎片化程度特征；

获取所述第二目标胚胎发育图像中的所有第二胚胎像素值；

对所有所述第二胚胎像素值进行遍历,获取第二极大像素值；

根据所述第二极大像素值的个数划分所述第二目标胚胎发育图像中的第二碎片化程

度特征。

6. 根据权利要求3所述的囊胚发育结果预测方法,其特征在于,所述通过具有所述训练细胞数量特征、所述训练均匀程度特征、所述训练碎片化程度特征以及所述训练组合特征的训练数据集对所述XGBoost预测模型进行训练,获得训练好的XGBoost预测模型,包括:

获取所述增强处理后的第一训练数据集中的所有所述第一胚胎发育图像中对应的每个细胞的第一训练轮廓,并根据所述每个细胞的第一训练轮廓,分割获得所有所述第一胚胎发育图像中对应的第一训练细胞数量特征;

获取所述增强处理后的第二训练数据集中的所有所述第二胚胎发育图像中对应的每个细胞的第二训练轮廓,并根据所述每个细胞的第二训练轮廓,分割获得所有所述第二胚胎发育图像中对应的第二训练细胞数量特征;

获取所有所述第一胚胎发育图像中对应的第一训练均匀程度特征;

获取所有所述第二胚胎发育图像中对应的第二训练均匀程度特征;

获取所述第一胚胎发育图像中的所有第一训练胚胎像素值;

对所有所述第一训练胚胎像素值进行遍历,获取第一极大训练像素值;

根据所述第一极大训练像素值的个数划分所述第一胚胎发育图像中的第一训练碎片化程度特征;

获取所述第二胚胎发育图像中的所有第二训练胚胎像素值;

对所有所述第二训练胚胎像素值进行遍历,获取第二极大训练像素值;

根据所述第二极大训练像素值的个数划分所述第二胚胎发育图像中的第二训练碎片化程度特征;

将所述第一训练细胞数量特征、所述第一训练均匀程度特征、所述第一训练碎片化程度特征、所述第二训练细胞数量特征、所述第二训练均匀程度特征和所述第二训练碎片化程度特征进行组合,获得训练组合特征;

通过具有所述第一训练细胞数量特征、所述第一训练均匀程度特征、所述第一训练碎片化程度特征、所述第二训练细胞数量特征、所述第二训练均匀程度特征、所述第二训练碎片化程度特征以及所述训练组合特征的所述训练数据集对所述XGBoost预测模型进行训练,获得训练好的XGBoost预测模型。

7. 根据权利要求5所述的囊胚发育结果预测方法,其特征在于,所述将具有所述目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序,包括:

将所述第一目标胚胎发育图像对应的所述第一细胞数量特征、所述第一均匀程度特征和所述第一碎片化程度特征和所述第二目标胚胎发育图像对应的所述第二细胞数量特征、所述第二均匀程度特征和所述第二碎片化程度特征进行组合,获得组合特征;

将所述第一细胞数量特征、所述第一均匀程度特征、所述第一碎片化程度特征、所述第二细胞数量特征、所述第二均匀程度特征、所述第二碎片化程度特征以及组合特征作为所述目标胚胎发育图像具有的特征;

将具有所述第一细胞数量特征、所述第一均匀程度特征、所述第一碎片化程度特征、所述第二细胞数量特征、所述第二均匀程度特征、所述第二碎片化程度特征以及组合特征的

所述目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序。

8. 一种囊胚发育结果预测系统,其特征在于,所述囊胚发育结果预测系统包括:

数据集获取单元,用于获取训练数据集,所述训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像;

特征获取单元,用于获取每张所述胚胎发育图像中的细胞数量特征、均匀程度特征和碎片化程度特征;

组合特征获取单元,用于将所述细胞数量特征、所述均匀程度特征和所述碎片化程度特征进行组合,获得组合特征;

模型构建单元,用于设置XGBoost模型中的参数,根据所述参数构建XGBoost预测模型;

模型训练单元,用于通过具有所述细胞数量特征、所述均匀程度特征、所述碎片化程度特征以及所述组合特征的训练数据集对所述XGBoost预测模型进行训练,获得训练好的XGBoost预测模型;

结果获取单元,用于获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,并将具有所述目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序。

9. 一种囊胚发育结果预测设备,其特征在于,包括至少一个控制处理器和用于与所述至少一个控制处理器通信连接的存储器;所述存储器存储有可被所述至少一个控制处理器执行的指令,所述指令被所述至少一个控制处理器执行,以使所述至少一个控制处理器能够执行如权利要求1至7任一项所述的囊胚发育结果预测方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机可执行指令,所述计算机可执行指令用于使计算机执行如权利要求1至7任一项所述的囊胚发育结果预测方法。

一种囊胚发育结果预测方法、系统、设备及存储介质

技术领域

[0001] 本发明涉及医疗人工智能技术领域,尤其是涉及一种囊胚发育结果预测方法、系统、设备及存储介质。

背景技术

[0002] 在生殖医学领域,囊胚培养本身是一个优胜劣汰的过程,能够培养成功的胚胎都是质量优异的,所以囊胚培养具有筛选优质胚胎的意义。但囊胚培养有一定的几率失败,而失败意味着浪费宝贵的胚胎资源。在这种情况下,如果能够在囊胚培养前对其发育结果进行预知,那么就能避免造成不必要的胚胎资源损失。

[0003] 临床上的传统方法。生殖科医生会依据自己的经验,在卵裂期(即选择囊胚发育之前的一个时期),依据一些指标(如胚胎分类的状态,是否均匀),对胚胎质量进行评估。对卵裂期胚胎质量评估是依据经验,对该阶段胚胎分裂数、均匀程度、碎片化程度来进行综合性评估。这个过程严重依赖医生的主观的经验性判断,具有不同临床工作经验的医生可能会给出不同结果,则对胚胎质量的判定结果也会不一致。而且该过程是人工操作,过程缓慢繁琐,不准确。

[0004] 而现有的深度学习技术分析囊胚发育延时影像没有实质解决胚胎预测问题。现有的深度学习技术的胚胎预测过程是基于囊胚发育期间的时间点,而不是在卵裂期的预测。并且现有方法对于卵裂期的图像分析的方法虽然能够对囊胚发育结果进行预测,但该过程不具备可解释性。而在临床上,对于结果的可解释性有一定的要求,这关系到患者对于结果的可信程度。

发明内容

[0005] 本发明旨在至少解决现有技术中存在的技术问题之一。为此,本发明提出一种囊胚发育结果预测方法、系统、设备及存储介质,能够使模型具有可解释性,解决医生对胚胎质量评估的主观性问题,提高预测囊胚发育结果的准确率。

[0006] 第一方面,本发明实施例提供了一种囊胚发育结果预测方法,所述囊胚发育结果预测方法包括:

获取训练数据集,所述训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像;

获取每张所述胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征;

将所述训练细胞数量特征、所述训练均匀程度特征和所述训练碎片化程度特征进行组合,获得训练组合特征;

设置XGBoost模型中的参数,根据所述参数构建XGBoost预测模型;

通过具有所述训练细胞数量特征、所述训练均匀程度特征、所述训练碎片化程度特征以及所述训练组合特征的训练数据集对所述XGBoost预测模型进行训练,获得训练好

的XGBoost预测模型；

获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征，并将具有所述目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测，获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序。

[0007] 与现有技术相比，本发明第一方面具有以下有益效果：

本方法为了能够解决医生对胚胎质量评估的主观性问题，提高预测囊胚发育结果的准确率，通过获取训练数据集，训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像；获取每张胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征；将训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征进行组合，获得训练组合特征；设置XGBoost模型中的参数，构建XGBoost预测模型；通过具有训练细胞数量特征、训练均匀程度特征、训练碎片化程度特征以及训练组合特征的训练数据集对XGBoost预测模型进行训练，获得训练好的XGBoost预测模型；获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征，并将具有目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测，获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。本方法获取每张胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征，能够解决医生对胚胎质量评估的主观性问题，通过将训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征进行组合，获得训练组合特征，能够提高预测囊胚发育结果的准确率，并通过XGBoost预测模型获得目标胚胎发育图像中对应特征的重要性排序，使得模型具有可解释性，使预测的囊胚发育结果具有可信度。

[0008] 根据本发明的一些实施例，所述获取训练数据集，所述训练数据集，包括：

获取在卵裂期的第一时间内进行囊胚培育并已知发育结果的多张第一胚胎发育图像，并将所有所述第一胚胎发育图像作为第一训练数据集；

获取在卵裂期的第二时间内进行囊胚培育并已知发育结果的多张第二胚胎发育图像，并将所有所述第二胚胎发育图像作为第二训练数据集；其中，所述第二时间与所述第一时间不同；

根据所述第一训练数据集和所述第二训练数据集生成训练数据集。

[0009] 根据本发明的一些实施例，所述目标胚胎发育图像包括第一目标胚胎发育图像和第二目标胚胎发育图像，所述第一目标胚胎发育图像和所述第二目标胚胎发育图像为目标胚胎对应的不同时间采集的胚胎发育图像，所述获取目标胚胎发育图像对应的细胞数量特征，包括：

采用Sobel算子和直方图均衡化对所述第一训练数据集中的每张所述第一胚胎发育图像进行图像增强处理，获得增强处理后的第一训练数据集；

在所述增强处理后的第一训练数据集中随机选取第一预设值的第一胚胎发育样本图像；

通过所述增强处理后的第一训练数据集对第一mask-RCNN实例分割模型进行分割训练，每次训练前对每张所述第一胚胎发育样本图像翻转一次，直到所述第一mask-RCNN实

例分割模型中的损失值收敛时结束训练,获得训练好的第一mask-RCNN实例分割模型;

通过所述训练好的第一mask-RCNN实例分割模型对所述第一目标胚胎发育图像进行实例分割,获得所述第一目标胚胎发育图像中的第一细胞数量特征和第一分割图像;

采用Sobel算子和直方图均衡化对所述第二训练数据集中的每张所述第二胚胎发育图像进行图像增强处理,获得增强处理后的第二训练数据集;

在所述增强处理后的第二训练数据集中随机选取第二预设值的第二胚胎发育样本图像;

通过所述增强处理后的第二训练数据集对第二mask-RCNN实例分割模型进行分割训练,每次训练前对每张所述第一胚胎发育样本图像翻转一次,直到所述第二mask-RCNN实例分割模型中的损失值收敛时结束训练,获得训练好的第二mask-RCNN实例分割模型;

通过所述训练好的第二mask-RCNN实例分割模型对所述第二目标胚胎发育图像进行实例分割,获得所述第二目标胚胎发育图像中的第二细胞数量特征和第二分割图像。

[0010] 根据本发明的一些实施例,所述获取目标胚胎发育图像对应的均匀程度特征,包括:

获取所述第一训练数据集中标记为均匀、不均匀和严重融合三类第一胚胎发育图像,并获取三类所述第一胚胎发育图像中对应的每个细胞的第一轮廓;

根据所述每个细胞的第一轮廓,分割获得三类所述第一胚胎发育图像中的多个细胞;

计算三类所述第一胚胎发育图像中对应的多个细胞中随机一个细胞的所有像素值,并计算所述细胞的所有像素值的第一中位数;

根据所述第一中位数,从三类所述第一胚胎发育图像中对应的所述第一分割图像中选取均匀、不均匀和严重融合对应的第一像素值范围;

计算所述第一目标胚胎发育图像对应的第一分割图像中所有细胞对应的第一目标中位数;

根据所述第一像素值范围和所述第一目标中位数,获取所述第一目标胚胎发育图像中的第一均匀程度特征;

获取所述第二训练数据集中标记为均匀、不均匀和严重融合三类第二胚胎发育图像,并获取三类所述第二胚胎发育图像中对应的每个细胞的第二轮廓;

根据所述每个细胞的第二轮廓,分割获得三类所述第二胚胎发育图像中的多个细胞;

计算三类所述第二胚胎发育图像中对应的多个细胞中随机一个细胞的所有像素值,并计算所述细胞的所有像素值的第二中位数;

根据所述第二中位数,从三类所述第二胚胎发育图像中对应的所述第二分割图像中选取均匀、不均匀和严重融合对应的第二像素值范围;

计算所述第二目标胚胎发育图像对应的第二分割图像中所有细胞对应的第二目标中位数;

根据所述第二像素值范围和所述第二目标中位数,获取所述第二目标胚胎发育图像中的第二均匀程度特征。

[0011] 根据本发明的一些实施例,所述获取目标胚胎发育图像对应的碎片化程度特征,

包括：

获取所述第一目标胚胎发育图像中的所有第一胚胎像素值；

对所有所述第一胚胎像素值进行遍历，获取第一极大像素值；

根据所述第一极大像素值的个数划分所述第一目标胚胎发育图像中的第一碎片化程度特征；

获取所述第二目标胚胎发育图像中的所有第二胚胎像素值；

对所有所述第二胚胎像素值进行遍历，获取第二极大像素值；

根据所述第二极大像素值的个数划分所述第二目标胚胎发育图像中的第二碎片化程度特征。

[0012] 根据本发明的一些实施例，所述通过具有所述训练细胞数量特征、所述训练均匀程度特征、所述训练碎片化程度特征以及所述训练组合特征的训练数据集对所述XGBoost预测模型进行训练，获得训练好的XGBoost预测模型，包括：

获取所述增强处理后的第一训练数据集中的所有所述第一胚胎发育图像中对应的每个细胞的第一训练轮廓，并根据所述每个细胞的第一训练轮廓，分割获得所有所述第一胚胎发育图像中对应的第一训练细胞数量特征；

获取所述增强处理后的第二训练数据集中的所有所述第二胚胎发育图像中对应的每个细胞的第二训练轮廓，并根据所述每个细胞的第二训练轮廓，分割获得所有所述第二胚胎发育图像中对应的第二训练细胞数量特征；

获取所有所述第一胚胎发育图像中对应的第一训练均匀程度特征；

获取所有所述第二胚胎发育图像中对应的第二训练均匀程度特征；

获取所述第一胚胎发育图像中的所有第一训练胚胎像素值；

对所有所述第一训练胚胎像素值进行遍历，获取第一极大训练像素值；

根据所述第一极大训练像素值的个数划分所述第一胚胎发育图像中的第一训练碎片化程度特征；

获取所述第二胚胎发育图像中的所有第二训练胚胎像素值；

对所有所述第二训练胚胎像素值进行遍历，获取第二极大训练像素值；

根据所述第二极大训练像素值的个数划分所述第二胚胎发育图像中的第二训练碎片化程度特征；

将所述第一训练细胞数量特征、所述第一训练均匀程度特征、所述第一训练碎片化程度特征、所述第二训练细胞数量特征、所述第二训练均匀程度特征和所述第二训练碎片化程度特征进行组合，获得训练组合特征；

通过具有所述第一训练细胞数量特征、所述第一训练均匀程度特征、所述第一训练碎片化程度特征、所述第二训练细胞数量特征、所述第二训练均匀程度特征、所述第二训练碎片化程度特征以及所述训练组合特征的所述训练数据集对所述XGBoost预测模型进行训练，获得训练好的XGBoost预测模型。

[0013] 根据本发明的一些实施例，所述将具有所述目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测，获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序，包括：

将所述第一目标胚胎发育图像对应的所述第一细胞数量特征、所述第一均匀程度特征和所述第一碎片化程度特征和所述第二目标胚胎发育图像对应的所述第二细胞数量特征、所述第二均匀程度特征和所述第二碎片化程度特征进行组合,获得组合特征;

将所述第一细胞数量特征、所述第一均匀程度特征、所述第一碎片化程度特征、所述第二细胞数量特征、所述第二均匀程度特征、所述第二碎片化程度特征以及组合特征作为所述目标胚胎发育图像具有的特征;

将具有所述第一细胞数量特征、所述第一均匀程度特征、所述第一碎片化程度特征、所述第二细胞数量特征、所述第二均匀程度特征、所述第二碎片化程度特征以及组合特征的所述目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序。

[0014] 第二方面,本发明实施例还提供了一种囊胚发育结果预测系统,所述囊胚发育结果预测系统包括:

数据集获取单元,用于获取训练数据集,所述训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像;

特征获取单元,用于获取每张所述胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征;

组合特征获取单元,用于将所述训练细胞数量特征、所述训练均匀程度特征和所述训练碎片化程度特征进行组合,获得训练组合特征;

模型构建单元,用于设置XGBoost模型中的参数,根据所述参数构建XGBoost预测模型;

模型训练单元,用于通过具有所述训练细胞数量特征、所述训练均匀程度特征、所述训练碎片化程度特征以及所述训练组合特征的训练数据集对所述XGBoost预测模型进行训练,获得训练好的XGBoost预测模型;

结果获取单元,用于获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,并将具有所述目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至所述训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和所述目标胚胎发育图像中对应特征的重要性排序。

[0015] 第三方面,本发明实施例还提供了一种囊胚发育结果预测设备,包括至少一个控制处理器和用于与所述至少一个控制处理器通信连接的存储器;所述存储器存储有可被所述至少一个控制处理器执行的指令,所述指令被所述至少一个控制处理器执行,以使所述至少一个控制处理器能够执行如上所述的一种囊胚发育结果预测方法。

[0016] 第四方面,本发明实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质存储有计算机可执行指令,所述计算机可执行指令用于使计算机执行如上所述的一种囊胚发育结果预测方法。

[0017] 可以理解的是,上述第二方面至第四方面与相关技术相比存在的有益效果与上述第一方面与相关技术相比存在的有益效果相同,可以参见上述第一方面中的相关描述,在此不再赘述。

附图说明

[0018] 本发明的上述和/或附加的方面和优点从结合下面附图对实施例的描述中将变得明显和容易理解,其中:

图1是本发明一实施例的一种囊胚发育结果预测方法的流程图;

图2是本发明另一实施例的一种囊胚发育结果预测方法的流程图;

图3是本发明一实施例的极大像素值示意图;

图4是本发明一实施例的一种囊胚发育结果预测系统的结构图。

具体实施方式

[0019] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,仅用于解释本发明,而不能理解为对本发明的限制。

[0020] 在本发明的描述中,如果有描述到第一、第二等只是用于区分技术特征为目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量或者隐含指明所指示的技术特征的先后关系。

[0021] 在本发明的描述中,需要理解的是,涉及到方位描述,例如上、下等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本发明和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。

[0022] 本发明的描述中,需要说明的是,除非另有明确的限定,设置、安装、连接等词语应做广义理解,所属技术领域技术人员可以结合技术方案的具体内容合理确定上述词语在本发明中的具体含义。

[0023] 在生殖医学领域,囊胚培养本身是一个优胜劣汰的过程,能够培养成功的胚胎都是质量优异的,所以囊胚培养具有筛选优质胚胎的意义。但囊胚培养有一定的几率失败,而失败意味着浪费宝贵的胚胎资源。在这种情况下,如果能够在囊胚培养前对其发育结果进行预知,那么就能避免造成不必要的胚胎资源损失。

[0024] 临床上的传统方法。生殖科医生会依据自己的经验,在卵裂期(即选择囊胚发育之前的一个时期),依据一些指标(如胚胎分类的状态,是否均匀),对胚胎质量进行评估。对卵裂期胚胎质量评估是依据经验,对该阶段胚胎分裂数、均匀程度、碎片化程度来进行综合性评估。这个过程严重依赖医生的主观的经验性判断,具有不同临床工作经验的医生可能会给出不同结果,则对胚胎质量的判定结果也会不一致。而且该过程是人工操作,过程缓慢繁琐,不准确。

[0025] 而现有的深度学习技术分析囊胚发育延时影像并没有实质解决预测问题。该过程基于囊胚发育期间的时间点,而不是在卵裂期的预测。对于卵裂期的图像分析的方法,分析模型能够对囊胚发育结果进行预测,但该过程是不具备可解释性的。而临床上,对于结果的可解释性有一定的要求,这关系到患者对于结果的可信程度。

[0026] 为解决上述问题,本发明为了能够解决医生对胚胎质量评估的主观性问题,提高预测囊胚发育结果的准确率,通过获取训练数据集,训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像;获取每张胚胎发育图像中的训练细胞数量特征、训

练均匀程度特征和训练碎片化程度特征;将训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征进行组合,获得训练组合特征;设置XGBoost模型中的参数,构建XGBoost预测模型;通过具有训练细胞数量特征、训练均匀程度特征、训练碎片化程度特征以及训练组合特征的训练数据集对XGBoost预测模型进行训练,获得训练好的XGBoost预测模型;获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,并将具有目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。本发明通过获取每张胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征,能够解决医生对胚胎质量评估的主观性问题,通过将训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征进行组合,获得训练组合特征,能够提高预测囊胚发育结果的准确率,并通过XGBoost预测模型获得目标胚胎发育图像中对应特征的重要性排序,使得模型具有可解释性,使预测的囊胚发育结果具有可信度。

[0027] 参照图1,本发明实施例提供了一种囊胚发育结果预测方法,本囊胚发育结果预测方法包括:

步骤S100、获取训练数据集,训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像。

[0028] 具体的,采用图像采集设备获取在卵裂期的第一时间内进行囊胚培育并已知发育结果的多张第一胚胎发育图像,并将所有第一胚胎发育图像作为第一训练数据集;

采用图像采集设备获取在卵裂期的第二时间内进行囊胚培育并已知发育结果的多张第二胚胎发育图像,并将所有第二胚胎发育图像作为第二训练数据集;其中,第二时间与第一时间不同;

根据第一训练数据集和第二训练数据集生成训练数据集。

[0029] 需要说明的是,本发明实施例中的图像采集设备可为公知的设备,本实施例不作具体限定。

[0030] 步骤S200、获取每张胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征。

[0031] 具体的,采用Sobel算子和直方图均衡化对第一训练数据集中的每张第一胚胎发育图像进行图像增强处理,通过人工标记获取增强处理后的第一训练数据集中的所有第一胚胎发育图像中对应的每个细胞的第一训练轮廓,并根据每个细胞的第一训练轮廓,分割获得所有第一胚胎发育图像中对应的第一训练细胞数量特征;

采用Sobel算子和直方图均衡化对第二训练数据集中的每张第二胚胎发育图像进行图像增强处理,通过人工标记获取增强处理后的第二训练数据集中的所有第二胚胎发育图像中对应的每个细胞的第二训练轮廓,并根据每个细胞的第二训练轮廓,分割获得所有第二胚胎发育图像中对应的第二训练细胞数量特征;

通过人工标记获取所有第一胚胎发育图像中对应的第一训练均匀程度特征;

通过人工标记获取所有第二胚胎发育图像中对应的第二训练均匀程度特征;

获取第一胚胎发育图像中的所有第一训练胚胎像素值;

对所有第一训练胚胎像素值进行遍历,获取第一极大训练像素值;

根据第一极大训练像素值的个数划分第一胚胎发育图像中的第一训练碎片化程度特征；

获取第二胚胎发育图像中的所有第二训练胚胎像素值；

对所有第二训练胚胎像素值进行遍历,获取第二极大训练像素值；

根据第二极大训练像素值的个数划分第二胚胎发育图像中的第二训练碎片化程度特征。

[0032] 在本实施例中,本发明通过图像处理与分析后提取每张胚胎发育图像中的细胞数量特征、均匀程度特征和碎片化程度特征,能够很好的解决传统人工胚胎质量评估受主观性影响的问题。

[0033] 步骤S300、将训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征进行组合,获得训练组合特征。

[0034] 具体的,将属于同一个胚胎的第一胚胎发育图像和第二胚胎发育图像合并为一个训练样本,将所有训练样本组成训练数据集,训练样本包括第一胚胎发育图像对应的第一细胞数量特征、第一均匀程度特征和第一碎片化程度特征,以及第二胚胎发育图像对应的第二细胞数量特征、第二均匀程度特征和第二碎片化程度特征；

将第一胚胎发育图像对应的第一训练细胞数量特征、第一训练均匀程度特征和第一训练碎片化程度特征,以及第二胚胎发育图像对应的第二训练细胞数量特征、第二训练均匀程度特征和第二训练碎片化程度特征进行组合,获得训练组合特征。

[0035] 步骤S400、设置XGBoost模型中的参数,根据参数构建XGBoost预测模型。

[0036] 具体的,设置XGBoost模型中的scale_pos_weight参数用于处理正样本和负样本分布不均衡的情况,正样本表示囊胚发育成功的样本,负样本表示囊胚发育失败的样本,若训练数据集中正样本数为PN,负样本数为NN,则设置scale_pos_weight = NN/PN;根据设置的scale_pos_weight参数,构建XGBoost预测模型。

[0037] 需要说明的是,本实施例中的XGBoost模型为公知常识,本实施例不作具体阐述,但需要通过更改XGBoost模型中的参数,以构建XGBoost预测模型。

[0038] 步骤S500、通过具有训练细胞数量特征、训练均匀程度特征、训练碎片化程度特征以及训练组合特征的训练数据集对XGBoost预测模型进行训练,获得训练好的XGBoost预测模型。

[0039] 具体的,通过具有第一训练细胞数量特征、第一训练均匀程度特征、第一训练碎片化程度特征、第二训练细胞数量特征、第二训练均匀程度特征、第二训练碎片化程度特征以及训练组合特征的训练数据集对XGBoost预测模型进行训练,获得训练好的XGBoost预测模型。

[0040] 步骤S600、获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,并将具有目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。

[0041] 具体的,通过图像采集设备获取目标胚胎发育图像,目标胚胎发育图像包括第一目标胚胎发育图像和第二目标胚胎发育图像,第一目标胚胎发育图像和第二目标胚胎发育图像为目标胚胎对应的不同时间采集的胚胎发育图像。获取第一目标胚胎发育图像和第二

目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,具体为:

获取目标胚胎发育图像对应的细胞数量特征包括:

采用Sobel算子和直方图均衡化对第一训练数据集中的每张第一胚胎发育图像进行图像增强处理,获得增强处理后的第一训练数据集;

在增强处理后的第一训练数据集中随机选取第一预设值的第一胚胎发育样本图像;

通过增强处理后的第一训练数据集对第一mask-RCNN实例分割模型进行分割训练,每次训练前对每张第一胚胎发育样本图像翻转一次,直到第一mask-RCNN实例分割模型中的损失值收敛时结束训练,获得训练好的第一mask-RCNN实例分割模型;

通过训练好的第一mask-RCNN实例分割模型对第一目标胚胎发育图像进行实例分割,获得第一目标胚胎发育图像中的第一细胞数量特征和第一分割图像;

采用Sobel算子和直方图均衡化对第二训练数据集中的每张第二胚胎发育图像进行图像增强处理,获得增强处理后的第二训练数据集;

在增强处理后的第二训练数据集中随机选取第二预设值的第二胚胎发育样本图像;

通过增强处理后的第二训练数据集对第二mask-RCNN实例分割模型进行分割训练,每次训练前对每张第一胚胎发育样本图像翻转一次,直到第二mask-RCNN实例分割模型中的损失值收敛时结束训练,获得训练好的第二mask-RCNN实例分割模型;

通过训练好的第二mask-RCNN实例分割模型对第二目标胚胎发育图像进行实例分割,获得第二目标胚胎发育图像中的第二细胞数量特征和第二分割图像。

[0042] 获取目标胚胎发育图像对应的均匀程度特征包括:

获取第一训练数据集中标记为均匀、不均匀和严重融合三类第一胚胎发育图像,并通过人工标记获取三类第一胚胎发育图像中对应的每个细胞的第一轮廓;

根据每个细胞的第一轮廓,分割获得三类第一胚胎发育图像中的多个细胞;

计算三类第一胚胎发育图像中对应的多个细胞中随机一个细胞的所有像素值,并计算细胞的所有像素值的第一中位数;

根据第一中位数,从三类第一胚胎发育图像中对应的第一分割图像中选取均匀、不均匀和严重融合对应的第一像素值范围;

计算第一目标胚胎发育图像对应的第一分割图像中所有细胞对应的第一目标中位数;

根据第一像素值范围和第一目标中位数,获取第一目标胚胎发育图像中的第一均匀程度特征;

获取第二训练数据集中标记为均匀、不均匀和严重融合三类第二胚胎发育图像,并通过人工标记获取三类第二胚胎发育图像中对应的每个细胞的第二轮廓;

根据每个细胞的第二轮廓,分割获得三类第二胚胎发育图像中的多个细胞;

计算三类第二胚胎发育图像中对应的多个细胞中随机一个细胞的所有像素值,并计算细胞的所有像素值的第二中位数;

根据第二中位数,从三类第二胚胎发育图像中对应的第二分割图像中选取均匀、

不均匀和严重融合对应的第二像素值范围；

计算第二目标胚胎发育图像对应的第二分割图像中所有细胞对应的第二目标中位数；

根据第二像素值范围和第二目标中位数，获取第二目标胚胎发育图像中的第二均匀程度特征。

[0043] 获取目标胚胎发育图像对应的碎片化程度特征包括：

获取第一目标胚胎发育图像中的所有第一胚胎像素值；

对所有第一胚胎像素值进行遍历，获取第一极大像素值；

根据第一极大像素值的个数划分第一目标胚胎发育图像中的第一碎片化程度特征；

获取第二目标胚胎发育图像中的所有第二胚胎像素值；

对所有第二胚胎像素值进行遍历，获取第二极大像素值；

根据第二极大像素值的个数划分第二目标胚胎发育图像中的第二碎片化程度特征。

[0044] 预测过程包括：

将第一目标胚胎发育图像对应的第一细胞数量特征、第一均匀程度特征和第一碎片化程度特征和第二目标胚胎发育图像对应的第二细胞数量特征、第二均匀程度特征和第二碎片化程度特征进行组合，获得组合特征；

将第一细胞数量特征、第一均匀程度特征、第一碎片化程度特征、第二细胞数量特征、第二均匀程度特征、第二碎片化程度特征以及组合特征作为目标胚胎发育图像具有的特征；

将具有第一细胞数量特征、第一均匀程度特征、第一碎片化程度特征、第二细胞数量特征、第二均匀程度特征、第二碎片化程度特征以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测，获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。

[0045] 需要说明的是，本发明实施例中的图像采集设备可为公知的设备，本实施例不作具体限定。

[0046] 需要说明的是，本实施例中的第一预设值和第二预设值可以根据实际需要进行修改，本实施例不作具体限定，以及本实施例中的根据第一极大像素值的个数划分第一碎片化程度特征和根据第二极大像素值的个数划分第二碎片化程度特征中第一极大像素值和第二极大像素值的个数也可以根据实际需要进行修改，本实施例不作具体限定。

[0047] 在本实施例中，通过图像处理与分析后获取每张胚胎发育图像中的训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征，能够解决医生对胚胎质量评估的主观性问题，通过将训练细胞数量特征、训练均匀程度特征和训练碎片化程度特征进行组合，获得训练组合特征，能够提高预测囊胚发育结果的准确率，并通过XGBoost预测模型获得目标胚胎发育图像中对应特征的重要性排序，使得模型具有可解释性，使预测的囊胚发育结果具有可信度。

[0048] 为了便于本领域人员理解，以下提供一组最佳实施例：

一、图像分析阶段：

参照图2,首先,针对在卵裂期(Day3)进行了囊胚培养并且已知发育结果的胚胎,例如,通过图像采集设备收集这些胚胎的第二天Day2的多张第一胚胎发育图像以及第三天Day3的多张第二胚胎发育图像,分别作为第一训练数据集dataset1和第二训练数据集dataset2,dataset1和dataset2训练数据集中图片数量均记为N。对于第一训练数据集dataset1和第二训练数据集dataset2,需预先标注每张胚胎发育图像的均匀程度。

[0049] 本实施例的目的是获得目标胚胎发育图像中的以下特征:

第一细胞数量特征D2num(胚胎发育第二天分裂的细胞数量),第一均匀程度特征D2uniformity(胚胎发育第二天的均匀程度),第一碎片化程度特征D2fragment(胚胎发育第二天的碎片化程度),第二细胞数量特征D3num(胚胎发育第三天分裂的细胞数量),第二均匀程度特征D3uniformity(胚胎发育第三天的均匀程度),第二碎片化程度特征D3fragment(胚胎发育第三天的碎片化程度)。

[0050] 需要说明的是,本发明实施例可以只获得卵裂期即Day3一天的胚胎图像进行分析,但获得Day2、Day3两天的胚胎图像进行分析,效果更好一下,准确率也更高一些。

[0051] 数据集的预处理:对于dataset1和dataset2中的所有胚胎图像,每张图像均裁减为312*312的宽高像素比,并且将裁剪后的图像转为灰度图(即单通道,每个像素值仅用一个数值表示,范围为0至255)。

[0052] 胚胎是位于培养皿中的,因此胚胎本身呈现圆状,将胚胎的最外围边缘,即培养皿的边框使用黑色(像素值设为R)进行绘制标记。

[0053] 在本实施例中,裁减为固定尺寸的目的:在图像处理时,方便对每一行列的像素值进行处理。边框绘制标记为黑色的目的:在图像处理时,遍历像素值,方便判定是否位于胚胎区域内部。

[0054] 通过以下方式提取目标胚胎发育图像中的各种特征:使用图像处理与分析实现胚胎分裂细胞个数的计数、胚胎均匀程度的分级、胚胎碎片化程度的分级。以下是提取细胞数量特征、均匀程度特征和碎片化程度特征的具体过程:

1、对胚胎分裂细胞个数进行计数:

通过mask-RCNN算法训练一个实例分割模型,获得第一mask-RCNN实例分割模型,通过第一训练数据集dataset1中的第一胚胎发育图像训练第一mask-RCNN实例分割模型,获得训练好的第一mask-RCNN实例分割模型。训练好的第一mask-RCNN实例分割模型能够将目标胚胎发育图像中分裂的各个细胞分割出来,然后统计数量,即得到分裂的细胞数量。

[0055] 第一步、图像预处理,具体为:

胚胎分裂细胞的状态是多个细胞可能在边缘部分有融合,图片存在区域模糊的情况,因此,对于图像分割具有一定的困难,且对于mask-RCNN算法来说,处理这类图像存在鲁棒性差的问题。

[0056] 对于此问题,本实施例的解决方法的先进行图像增强处理,让胚胎图像中各个小细胞的边缘部分更加凸显,增强不同实例的像素差异,以减小实例分割的难度、增加分割的准确度,使结果更加客观。

[0057] 具体地,使用Sobel算子对dataset1中的每张第一胚胎发育图像的分裂细胞的轮廓进行锐化处理,并使用直方图均衡化进一步增强分裂细胞和背景的对比度,这样一个过程,能够提高内部小细胞的边缘和背景的分度,减小后续的实例分割任务难度。

[0058] 处理完后,通过人工标记对边缘轮廓进行标记。

[0059] 第二步、第一mask-RCNN实例分割模型的训练过程:

将dataset1用于第一mask-RCNN实例分割模型的训练,直到第一mask-RCNN实例分割模型中的损失值收敛时(即损失值不再变小时)结束训练,获得训练好的第一mask-RCNN实例分割模型。考虑到胚胎图像中细胞融合的区域不同,可能会影响模型分割的准确度。本实施例选择在每一次训练的开始时,随机选取数据集中50%数量的样本图像进行翻转操作(翻转角度选取 1° 至 360° 的随机值),这样能够提高图像分割模型的泛化能力。

[0060] 第三步、对目标胚胎发育图像进行实例分割,具体为:

通过训练好的第一mask-RCNN实例分割模型对第一目标胚胎发育图像进行实例分割,获得第一目标胚胎发育图像中的第一细胞数量特征和第一分割图像。在图像分割的过程中,对每一个实例进行序数标记,用于细胞数量的计数,即每次分割得到一个实例,则累计加一,每次分割完毕输出计数值。

[0061] 需要说明的是,由于通过第一训练数据集dataset2中的第二胚胎发育图像训练第二mask-RCNN实例分割模型,获得训练好的第二mask-RCNN实例分割模型的步骤过程与获得训练好的第一mask-RCNN实例分割模型的步骤过程一样,本实施例不作详细描述,因此,通过训练好的第二mask-RCNN实例分割模型对第二目标胚胎发育图像进行实例分割,获得第二目标胚胎发育图像中的第二细胞数量特征和第二分割图像。

[0062] 2、对胚胎均匀程度进行分级:

本实施例将所有第一胚胎发育图像中所表现的均匀程度定义为三种类别:均匀,不均匀,严重融合。即,本实施例的目标就是将一张目标胚胎发育图像分类为这三种类别之一。

[0063] 在均匀情况下,胚胎分裂的每一个小细胞的轮廓都清晰可见,而且整个小细胞内部趋于平滑,很容易区分是一个独立的小细胞;在不均匀的情况下,对于一个胚胎,整体上一般表现为某一小块不均匀,例如,某个角上的2个小细胞融合在了一起;在严重融合的情况下,相对于“不均匀”的情况,此时应该是大部分区域都是细胞融合的状态,而不仅仅是一小块区域有融合是情况。目前有些研究可能使用图像识别的方法,将均匀,不均匀,严重融合,定义为一个三分类问题。仅仅使用图像识别去对这三种类别进行分类,实现效果可能不理想。例如:

对于一张胚胎图像X1,假设它靠左上角区域是融合的,则必定将此定义为“不均匀”类别。又有另一张胚胎图像X2,其右下角区域是融合的,X2也应该被评估为“不均匀”。但是在这种情况下,由于图像识别的本质是提取图像的特征,特征是用于区分不同类别的主要因素,所以,对于X1,它的主要区分特征位于左上角,而X2主要用于区分类别的特征位于右下角,此时,图像识别期就不会具有很大概率把X1和X2归为一类。但事实是,他们都应该被归为“不均匀”类别。

[0064] 所以,这就是图像识别方法在此处不适用的原因,因为胚胎图像变化情况多,对于类别之间的界定是难以清晰地表达的。

[0065] 本实施例解决此问题的方法是:使用基于细胞单区域内像素区间划分的方法,如果一个胚胎内部细胞是均匀的,则单个细胞内的所有像素值是趋于单一的,单个细胞内的所有像素值不一定完全相等,但是必定位于一个数值范围内。而对于“不均匀”的胚胎,各小

细胞内部颜色会稍深(因为存在融合,重叠),因此,像素值所处于的数值范围相比较于“均匀”的胚胎会有所区别。对于“严重融合”的胚胎,同样如此,细胞内部颜色会更深,像素值所处于数值区间也会和前两者有所差别。

[0066] 相对于图像识别方法,本实施例中的方法是聚焦于单个细胞,这样做能考虑到整张胚胎图像中各个细胞的单独情况,最后综合考虑得出结果,而不会造成像图像识别会因为单独一小块区域而对模型结果造成较大扰动情况。

[0067] 本实施例中的方法执行过程具体如下:

以下步骤是在利用第一mask-RCNN实例分割模型得到目标胚胎发育图像对应的第一分割图像的各个细胞的基础之下进行的操作。

[0068] 第一步:计算不同分类的像素值正常范围,具体为:

本实施例先通过人工标记对dataset1中所有第一胚胎发育图像进行“均匀”,“不均匀”,“严重融合”三类均匀程度特征的标记,分别把这3类均匀程度特征记为EVEN_IMG类,UNEVEN_IMG类,FUS_IMG类。每一类对应的第一胚胎发育图像的数量分别为EVEN_IMG_N,UNEVEN_IMG_N,FUS_IMG_N。

[0069] A. 对于EVEN_IMG类:

EVEN_IMG_N张第一胚胎发育图像都是“均匀”类别的。通过人工标记获取“均匀”类别的第一胚胎发育图像中对应的每个细胞的第一轮廓,根据每个细胞的第一轮廓,分割获得“均匀”类别的第一胚胎发育图像中的多个细胞,对于每张“均匀”的第一胚胎发育图像,随机选择一个单独的细胞,统计细胞内部像素值情况,计算第一中位数。这样,就有EVEN_IMG_N个第一中位数,对于这些第一中位数:

本实施例选取 $[Q1-3*IQR, Q3+3*IQR]$,其中, $[]$ 表示闭区间,作为正常像素值的范围,其中Q1表示所有像素值的25%分位数,Q3表示所有像素值的75%分位数, $IQR=Q3-Q1$,表示四分位距。这个过程就相当于,把位于 $[Q1-3*IQR, Q3+3*IQR]$ 内部的像素值作为“均匀”的参考标准,因为这个范围代表了绝大多数“均匀”细胞的像素值范围,所以它具有代表性;

B. 对于UNEVEN_IMG类,处理方法与EVEN_IMG类相同;

C. 对于FUS_IMG类,处理方法与EVEN_IMG类相同;

D. 至此,就能分别得到三类第一胚胎发育图像的像素值正常范围,也就是有三组 $[Q1-3*IQR, Q3+3*IQR]$,分别记为均匀像素范围EVEN_RANGE、不均匀像素范围UNEVEN_RANGE、严重融合像素范围FUS_RANGE。

[0070] 第二步:提取均匀程度特征,具体为:

对于第一目标胚胎发育图像,本实施例想划定它是属于“均匀”,“不均匀”和“严重融合”这三类均匀程度特征中的哪一种,则在通过第一mask-RCNN实例分割模型得到每个小细胞的编号的情况下,遍历每一个小细胞,统计所有小细胞的像素值情况(即,计算第一目标中位数),对这些像素值落入上述三个区间内的数量进行计数,如果某个像素值落入多个区间(区间可能有像素值重叠),则将落入多个区间的数量都进行累计,最后得到第一胚胎发育图像的所有细胞的像素值落入三个区间内的累计数,分别记为CNT_EVEN,CNT_UNEVEN,CNT_FUS,最后取三者的最大值,即,将累计数最大的区间对应类别作为第一目标胚胎发育图像的均匀程度分级结果,即,第一目标胚胎发育图像的均匀程度特征。

[0071] 需要说明的是,本实施例通过人工标记获取所有第一胚胎发育图像中对应的第一训练均匀程度特征;通过人工标记获取所有第二胚胎发育图像中对应的第二训练均匀程度特征。由于采用第二训练数据集dataset2获得第二目标胚胎发育图像的第二均匀程度特征与采用第一训练数据集dataset1获得第一目标胚胎发育图像的第一均匀程度特征的方式一样,本实施例不作详细描述,因此,基于第二训练数据集dataset2,能够获得第二目标胚胎发育图像的第二均匀程度特征。

[0072] 本实施例通过图像处理的方式获取细胞的中位数,并计算获得三类均匀程度特征对应的像素范围值,能够解决医生对胚胎质量评估的主观性问题。

[0073] 3、对胚胎碎片化程度进行分级:

胚胎发育的图像当中,如果胚胎分裂的细胞数很清晰,而且比较均匀的话,则表面会表现的很平滑(呈现一个个细胞的状态),而如果胚胎趋于融合的状态,细胞之间的边缘会互相趋于一体,这种状态下,在融合处就会显得凹凸不平,呈现多种碎片状态。因此,如果第一目标胚胎发育图像表现的凹凸不平的程度越大,说明胚胎的碎片化程度越高。

[0074] 对胚胎碎片化程度的分级定义为:A、B、C级,其中,A为碎片化程度最低,C为碎片化程度最高。

[0075] 本实施例提出了一种使用像素的峰值对胚胎碎片化程度进行评估的方法,称为:基于像素的峰值计数。基于像素的峰值计数方法具体如下(以下是对一张图的处理,数据集中所有图像的碎片化分级方法同理):

胚胎发育图像的像素凹凸程度决定了碎片化程度,平滑是与凹凸不平相反的状态。平滑意味着该胚胎表面各区域接收光的角度差异不大,则该表面的像素值差异较小。相比较之下,如果一个胚胎表面是凹凸不平的,则凹陷处和凸显处像素差值会较大,这是从接受光的角度分析。与此同时,胚胎中分裂的细胞在融合处(即凹陷处)也会与细胞表面颜色深浅差异较大。因此,基于以上几点,参照图3,图3为极大像素值示意图,其中,(a)图为像素表面平滑的极大像素值示意图,(b)图为像素表面凹凸不平的极大像素值示意图。(a)图的碎片化程度低,则胚胎发育图像的像素表面平滑,像素极大值点少,(b)图的碎片化程度高,则胚胎发育图像的像素表面凹凸不平,像素极大值点多,图中黑点表示像素的峰值(即,像素极大值),基于像素的峰值计数方法的具体流程如下:

第一步、对312*312的第一目标胚胎发育图像,依据像素值所在列,逐一按照从上到下、从左往右的顺序进行遍历,即有312列,每列有312个像素值。直到第一次遇到像素值R(即为开始标记的胚胎区域的边框),则开始寻找像素的极大值,遍历到第二次遇到R,说明该列遍历结束,进行下一列的遍历。

[0076] 第二步、该步骤是为了找该列中的像素的极大值,即,获得第一极大像素值,一个极大值表示一个峰值,每一个峰值能够代表一个凸表面,累加峰值的个数。

[0077] 第三步、对所有的列重复第二步,得到第一目标胚胎发育图像的像素峰值总数,即,获得第一极大像素值的个数。

[0078] 得到了第一目标胚胎发育图像的像素峰值总数(即,获得第一极大像素值的个数)后,依据像素峰值总数进行分级,获得碎片化程度分级结果,即,第一目标胚胎发育图像的第一碎片化程度特征。定义碎片化程度分级为:

峰值数小于或等于10时为碎片化程度A级;

峰值数处于10至50时为碎片化程度B级；

峰值数大于50时为碎片化程度C级。

[0079] 需要说明的是,由于第二训练数据集dataset2中的第二训练碎片化程度特征的获取方式、第一训练数据集dataset1中的第一训练碎片化程度特征的获取方式以及第二目标胚胎发育图像中的第二碎片化程度特征获取方式,与第一目标胚胎发育图像中的第一碎片化程度特征的获取方式一样,本实施例不作详细描述。

[0080] 至此,在图像分析阶段,对第一目标胚胎发育图像、第二目标胚胎发育图像、第一训练数据集dataset1以及第二训练数据集dataset2的处理已经完毕,这样就能够获得第一目标胚胎发育图像的D2num、D2uniformity、D2fragment 这三个值的客观分析结果,获得第二目标胚胎发育图像的D3num、D3uniformity、D3fragment的值,获得第一训练数据集dataset1的第一训练细胞数量特征、第一训练均匀程度特征、第一训练碎片化程度特征,获得第二训练数据集dataset2的第二训练细胞数量特征、第二训练均匀程度特征、第二训练碎片化程度特征。

[0081] 二、文本数据分析阶段:

在获得第一训练细胞数量特征、第一训练均匀程度特征、第一训练碎片化程度特征、第二训练细胞数量特征、第二训练均匀程度特征和第二训练碎片化程度特征这六个特征对应的值后,构建一个样本数量为N、特征数量为六的训练数据集dataset,即,将属于同一个胚胎的第一胚胎发育图像和第二胚胎发育图像合并为一个训练样本,将所有训练样本组成训练数据集dataset,其中,训练样本包括第一胚胎发育图像对应的第一训练细胞数量特征D2num2、第一训练均匀程度特征D2uniformity2和第一训练碎片化程度特征D2fragment2,以及第二胚胎发育图像对应的第二训练细胞数量特征D3num3、第二训练均匀程度特征D3uniformity3和第二训练碎片化程度特征D3fragment3。

[0082] 在这一部分,根据样本数据集dataset中的D2num2、D2uniformity2、D2fragment2、D3num3、D3uniformity3、D3fragment3这六个特征,进行机器学习分类器的构建。具体为:

首先,考虑到胚胎发育数据的时序影响(即第二天和第三天前后的特征存在一定的关联性),本实施例将D2num2、D2uniformity2、D2fragment2、D3num3、D3uniformity3、D3fragment3这六个独立特征进行特征组合,生成新的特征,即,将第二天和第三天的特征分别进行前后顺序的组合,则总组合数为九组,组合过程为:D2num2和D3num3组合,D2num2和 D3uniformity3组合,D2num2和D3fragment3组合,D2uniformity2和D3num3组合,D2uniformity2和 D3uniformity3组合,D2uniformity2和D3fragment3组合,D2fragment2和D3num3组合,D2fragment2和 D3uniformity3组合,D2fragment2和D3fragment3组合。因此,特征总数为15,即,六个独立特征和九个组合特征。

[0083] 使用XGBoost模型构建囊胚培养发育结果预测模型,即,XGBoost预测模型。数据集dataset中可能存在样本分布不平衡的问题,因此,将XGBoost模型的参数设置如下:

定义囊胚发育成功的样本为正样本,囊胚发育失败的样本为负样本,scale_pos_weight参数用于处理正样本和负样本分布不平衡的情况,设数据集dataset中正样本数为PN,负样本数为NN,则设置scale_pos_weight = NN/PN,scale_pos_weight参数值越大,则XGBoost预测模型中对于少数类样本的学习率就会增大,相当于增加了少数类样本的重要性。

[0084] 将训练数据集dataset中的六个独立特征和九个组合特征输入至XGBoost预测模型中进行训练,在XGBoost预测模型收敛后结束训练,获得训练好的XGBoost预测模型。

[0085] 将训练好的XGBoost预测模型直接用于目标胚胎发育图像的预测,过程为:

将第一目标胚胎发育图像对应的第一细胞数量特征D2num、第一均匀程度特征D2uniformity和第一碎片化程度特征D2fragment和第二目标胚胎发育图像对应的第二细胞数量特征D3num、第二均匀程度特征D3uniformity和第二碎片化程度特征D3fragment进行组合,获得组合特征;

将第一细胞数量特征D2num、第一均匀程度特征D2uniformity和第一碎片化程度特征D2fragment、第二细胞数量特征D3num、第二均匀程度特征D3uniformity和第二碎片化程度特征D3fragment以及组合特征作为目标胚胎发育图像具有的特征;

将具有第一细胞数量特征D2num、第一均匀程度特征D2uniformity和第一碎片化程度特征D2fragment、第二细胞数量特征D3num、第二均匀程度特征D3uniformity和第二碎片化程度特征D3fragment以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。

[0086] 通过训练好的XGBoost预测模型输出目标胚胎发育图像对应的feature_importance列表,该列表输出为15个特征(即,六个独立特征对应的输出和九个组合特征对应的输出)的重要性排序,因此,能够根据排序结果对模型进行可解释性分析,即哪些因素对囊胚发育的结果具有重要性影响,而哪些因素影响程度较小。

[0087] 在本实施例中,由于初始的六个特征被认为是独立不相关的,而实际情况是他们之间存在时间上的组合,即前者可能影响后者,因此,将组合特征用于模型训练,相当于对时序特征进行了挖掘,而时序特征可能对模型具有影响。因此,通过将六个独立特征和九个组合特征输入至XGBoost预测模型进行训练,获得训练好的XGBoost预测模型;通过训练好的XGBoost预测模型对目标胚胎发育图像进行预测,能够解决医生对胚胎质量评估的主观性问题,提高预测囊胚发育结果的准确率,并通过XGBoost预测模型对目标胚胎发育图像进行预测,获得目标胚胎发育图像中对应特征的重要性排序,使得模型具有可解释性,使预测的囊胚发育结果具有可信度。

[0088] 参照图4,本发明实施例还提供了一种囊胚发育结果预测系统,本囊胚发育结果预测系统包括数据集获取单元100、特征获取单元200、组合特征获取单元300、模型构建单元400、模型训练单元500和结果获取单元600,其中:

数据集获取单元100,用于获取训练数据集,训练数据集包括卵裂期进行囊胚培育并已知发育结果的多张胚胎发育图像;

特征获取单元200,用于获取每张胚胎发育图像中的细胞数量特征、均匀程度特征和碎片化程度特征;

组合特征获取单元300,用于将细胞数量特征、均匀程度特征和碎片化程度特征进行组合,获得组合特征;

模型构建单元400,用于设置XGBoost模型中的参数,构建XGBoost预测模型;

模型训练单元500,用于通过具有细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的训练数据集对XGBoost预测模型进行训练,获得训练好的XGBoost预测模

型；

结果获取单元600,用于获取目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征,并将具有目标胚胎发育图像对应的细胞数量特征、均匀程度特征、碎片化程度特征以及组合特征的目标胚胎发育图像输入至训练好的XGBoost预测模型中进行预测,获得预测的囊胚发育结果和目标胚胎发育图像中对应特征的重要性排序。

[0089] 需要说明的是,由于本实施例中的一种囊胚发育结果预测系统与上述的一种囊胚发育结果预测方法基于相同的发明构思,因此,方法实施例中的相应内容同样适用于本系统实施例,此处不再详述。

[0090] 本发明实施例还提供了一种囊胚发育结果预测设备,包括:至少一个控制处理器和用于与至少一个控制处理器通信连接的存储器。

[0091] 存储器作为一种非暂态计算机可读存储介质,可用于存储非暂态软件程序以及非暂态性计算机可执行程序。此外,存储器可以包括高速随机存取存储器,还可以包括非暂态存储器,例如至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施方式中,存储器可选包括相对于处理器远程设置的存储器,这些远程存储器可以通过网络连接至该处理器。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0092] 实现上述实施例的一种囊胚发育结果预测方法所需的非暂态软件程序以及指令存储在存储器中,当被处理器执行时,执行上述实施例中的一种囊胚发育结果预测方法,例如,执行以上描述的图1中的方法步骤S100至步骤S600。

[0093] 以上所描述的系统实施例仅仅是示意性的,其中作为分离部件说明的单元可以是或者也可以不是物理上分开的,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0094] 本发明实施例还提供了一种计算机可读存储介质,该计算机可读存储介质存储有计算机可执行指令,该计算机可执行指令被一个或多个控制处理器执行,可使得上述一个或多个控制处理器执行上述方法实施例中的一种囊胚发育结果预测方法,例如,执行以上描述的图1中的方法步骤S100至步骤S600的功能。

[0095] 本领域普通技术人员可以理解,上文中所公开方法中的全部或某些步骤、系统可以被实施为软件、固件、硬件及其适当的组合。某些物理组件或所有物理组件可以被实施为由处理器,如中央处理器、数字信号处理器或微处理器执行的软件,或者被实施为硬件,或者被实施为集成电路,如专用集成电路。这样的软件可以分布在计算机可读介质上,计算机可读介质可以包括计算机存储介质(或非暂时性介质)和通信介质(或暂时性介质)。如本领域普通技术人员公知的,术语计算机存储介质包括在用于存储信息(诸如计算机可读指令、数据结构、程序模块或其他数据)的任何方法或技术中实施的易失性和非易失性、可移除和不可移除介质。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字多功能盘(DVD)或其他光盘存储、磁盒、磁带、磁盘存储或其他磁存储装置、或者可以用于存储期望的信息并且可以被计算机访问的任何其他的介质。此外,本领域普通技术人员公知的是,通信介质通常包含计算机可读指令、数据结构、程序模块或者诸如载波或其他传输机制之类的调制数据信号中的其他数据,并且可包括任何信息递送介质。

[0096] 上面结合附图对本发明实施例作了详细说明,但本发明不限于上述实施例,在所属技术领域普通技术人员所具备的知识范围内,还可以在不脱离本发明宗旨的前提下作出各种变化。

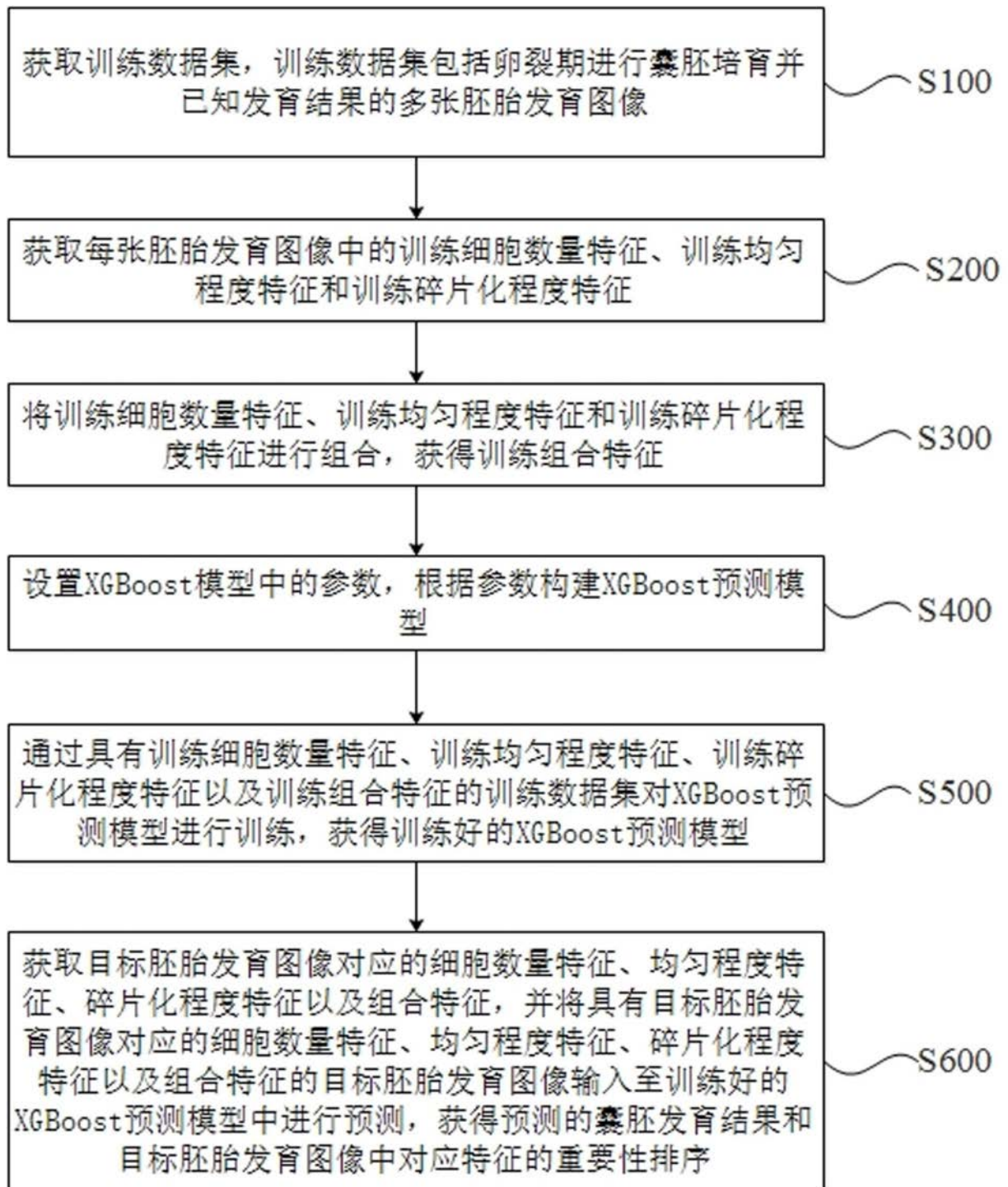


图1

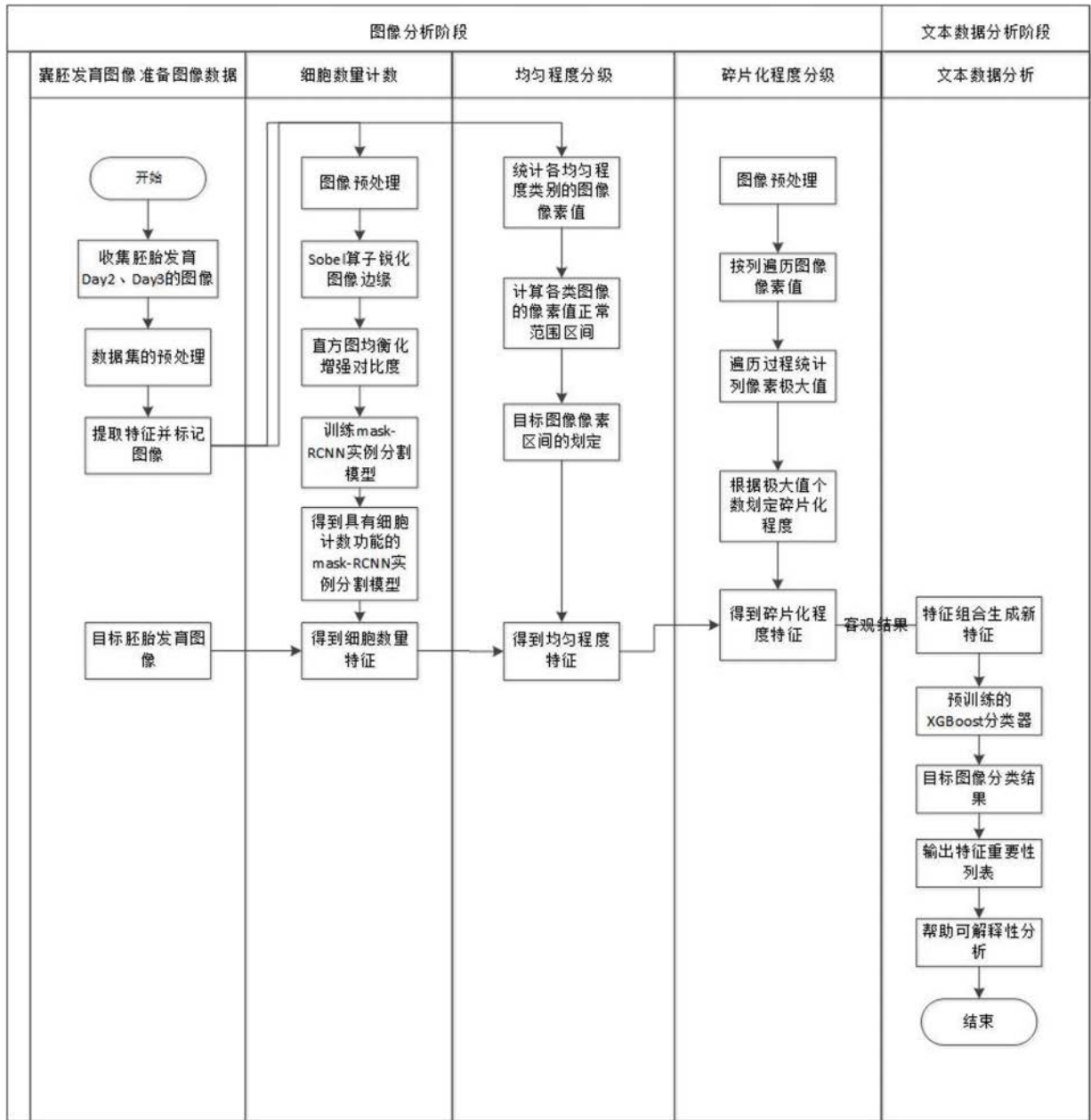


图2

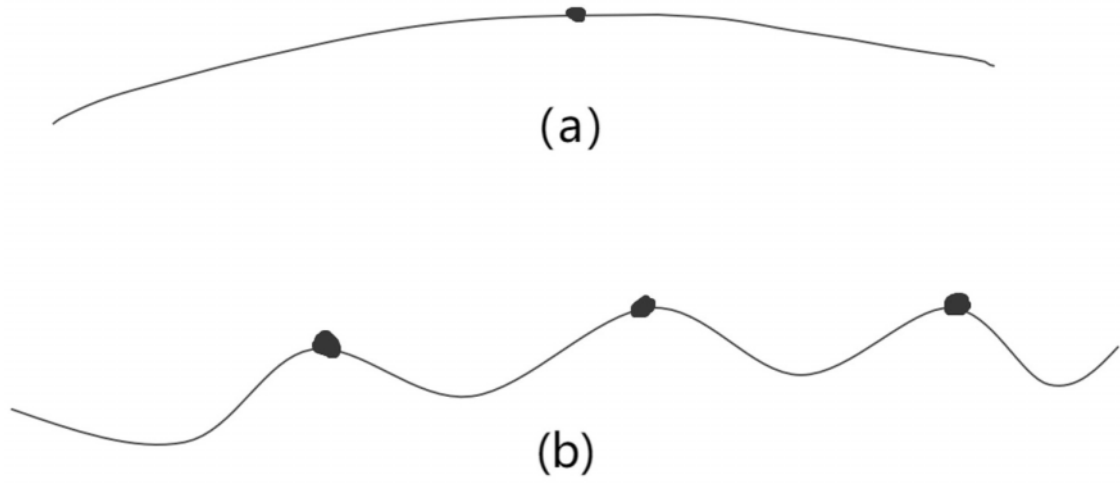


图3

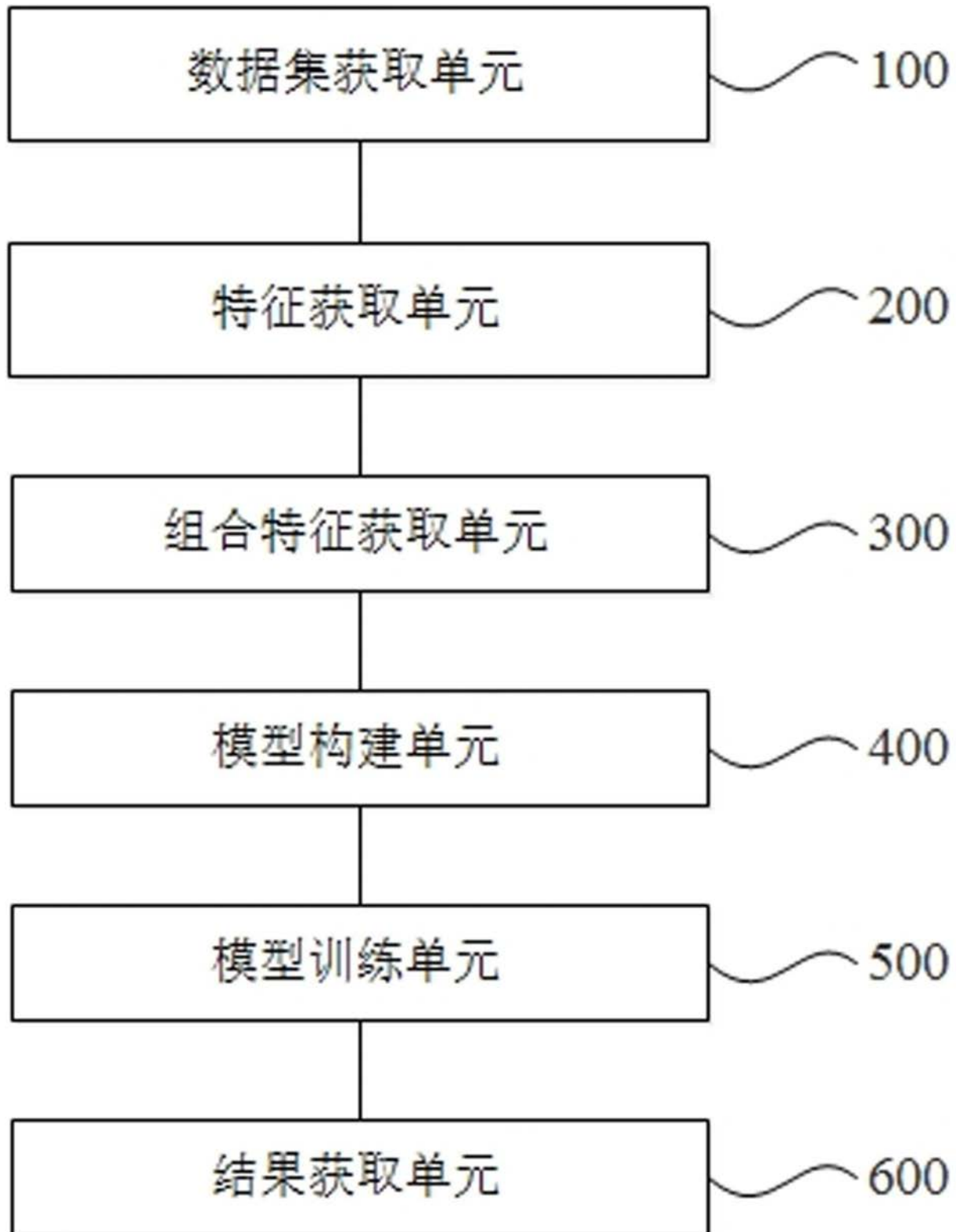


图4